Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 3 (2014) 962 – 971

Transportation Research Procedia

www.elsevier.com/locate/procedia

17th Meeting of the EURO Working Group on Transportation, EWGT2014, 2-4 July 2014, Sevilla, Spain

Calibration of microsimulation models – The effect of calibration parameters errors in the models' performance

Miguel Figueiredo[a], Álvaro Seco[a], Ana Bastos Silva[a*]

[a]*CITTA, Department of Civil Engineering, University of Coimbra, Rua Luís Reis Santos - Pólo II da Universidade, 3030-788 Coimbra, Portugal*

Abstract

Microsimulation models are now widely used. However, the capacity of a model to represent the reality with high level of accuracy depends significantly on the calibration process. This paper presents a sensitivity analysis to test how the expected accuracy of the traffic microsimulation models' outputs can be affected by different errors' types and degrees in the estimation of calibration parameters.

Using Aimsun software it was possible to establish different relations between the level of the calibration parameters errors, and the corresponding errors in the simulated results. The significant importance of the "reaction time" and "minimum distance between vehicles" was confirmed.

Selection and peer-review under responsibility of the Scientific Committee of EWGT2014

Keywords: microsimulation; AIMSUN; calibration; calibration parameters; errors; accuracy

1. Introduction

Nowadays microscopic traffic simulation models are widely used in several stages of project and with different objectives. These models represent the reality with high detail relatively to the infrastructure network, traffic demand, drivers' behavior, vehicles dynamics and route choice. They make use of a wide range of information which needs to be collected in the field, and are dependent on a significant number of calibration parameters, which are subject to several potential estimation errors and which can significantly affect the models' forecasting ability. In fact, in many cases, these parameters prove difficult to accurately quantify because: some are difficult to collect and

* Corresponding author. *E-mail address:* abastos@dec.uc.pt

Selection and peer-review under responsibility of the Scientific Committee of EWGT2014
doi:10.1016/j.trpro.2014.10.076

demand expensive and inaccessible equipment; there are imprecisions in the data collected; unavailability of field data forces the adoption of estimates produced by other models.

In the coding/construction of a model the calibration stage is especially important to assure that the model represents the reality in an accurate way. The calibration process is, usually, time-consuming and dependent on the quantity and quality of the information to be used. Thus, an important modelling question is where, within the calibration process, is it more important to apply the resources available?

The FHWA (2004) refers that it is of equal importance to define what is going to be studied as is what is not. Also, it is necessary to define the precision (minimum and maximum) levels that it is necessary to associate to each parameter, to ensure that, in a cost-effective way, the model represents the reality satisfactorily for the specific purpose that it is being built, since the cost of developing a model increases rapidly when high levels of trust are required (Vasconcelos et al, 2009; Sargent, 2000). In fact, trying to reduce to zero the quantification errors of parameters, is not justified for the majority of parameters, either because that is not achievable, or because the marginal costs associated to it are superior to the marginal benefit resulting from the increase in the quality of the models results.

To support this kind of calibration decisions, it is especially important to better understand and quantify the relations between the input parameters' errors and the models' output precision levels that result from their usage.

2. Methodological Approach

To develop this study five basic and sequential steps were followed: (i) detailed study of the microscopic models which are the base of the microsimulator software; (ii) listing the calibration parameters involved in the different microscopic models, their meaning and expected type of influence in the model outputs; (iii) selection of calibration parameters to study; (iv) selection of performance indicators to use in analyzing results; and (v) evaluation of the impacts caused by error introduction in the performance of the model, e.g., quality of results. To develop this study, the AIMSUN microsimulation software was used (version 7.0.4).

Within this framework, steps i and ii were based in the micro models directly embed in this software, as well in their theoretical approaches. It was opted to develop all the study based on the construction/coding of a model of one intersection, entirely regulated by traffic lights, based on a real case. This option enabled to support the model coding by site data collection, essential to the coding, calibration (non-exhaustive), and validation procedures. In the analysis and evaluation of results it was assumed that the reference coded model was representative of "one possible reality", and not obligatorily of "one observed reality".

Taking into account the case study characteristics and, therefore, the parameters in which an introduction of errors will lead to bigger impacts in the model results, the selection of parameters to study was made (step iii).

Step iv focused in the selection of performance indicators. The AIMSUN software provides a large number of possible indicators (traffic flows, delays, travel times, etc), which can represent global network results or individual or partial network ones (per vehicle, certain route, infrastructural element, etc.). A robust performance indicator was pursued, capable of representing the global behavior of the model and be sensible enough to error introduction in the selected calibration parameters. For simplicity of the analyses, it was opted to select only one performance indicator. The "average travel time" was chosen since it closely correlates with the performance of an intersection and with its degree of saturation.

In the last work phase (step v) a systematic evaluation of the impact on the outputs caused by input errors in calibration parameters is conducted.

3. Case study

For the development of this work an appropriate case study was selected to develop a model of reference. Reference microsimulation guidelines (FHWA, 2004; Austroads, 2006) and studies (Punzo & Ciuffo, 2009, Hollander & Liu, 2008), that specifically make an approach to this matter, mainly present freeways, or networks in rural environments as case studies. It was believed that a smaller network, within an urban environment could be a relevant area of study. For this reason, it was chosen a single but complex signalized intersection.

The selected case study was the complex and busy intersection *Arnado – Auto Industrial*, fully controlled by traffic lights, in the city of Coimbra's downtown. A network scheme is presented in Figure 1. There were made traffic counts in the period of 7h30 to 9h30 of the morning, in a regular working day. There were counted a total of 3970 vehicles, of which 3574 light vehicles. An AIMSUN model was coded taking into account all the field data. As the "reference model" was intended to represent only "one possible reality", little effort to calibrate the model was made. Instead, it was decided to maintain all the AIMSUN default parameter values enabling a more useful practical interpretation of the results, since they would result from errors within typical ranges of model parameters values.

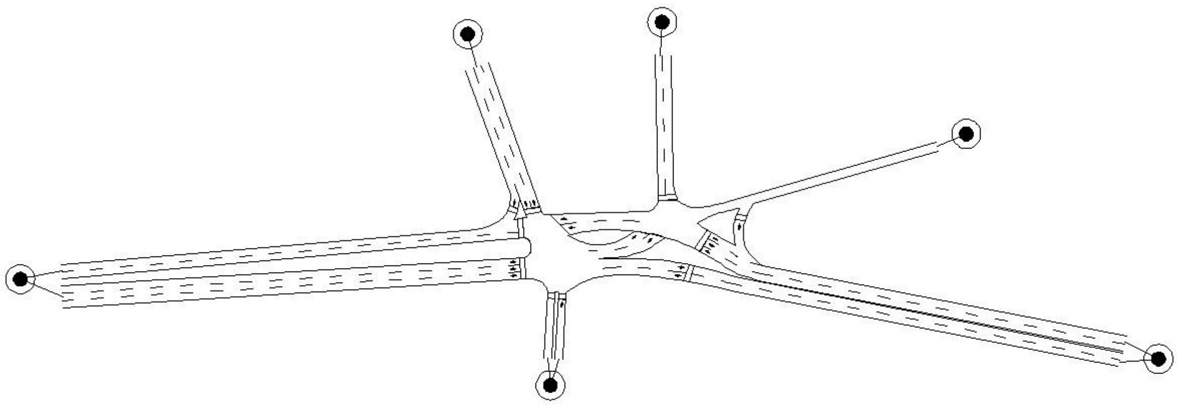


Fig. 1. Schematic view of the case study: a busy urban intersection.

4. Parameters selection

The values adopted in the wide range of calibration parameters incorporated in transport microsimulation models and sub-models, can significantly influence the modeled outputs. They vary from those relative to the characteristics and behavior of drivers, to those relative to the vehicles dynamics and to infrastructures' basic characteristics. Parameters can also be relative to driver route choice behavior.

In the AIMSUN software there is a categorization of calibration parameters according to three big groups, independently of the models that they affect: global, local or vehicle type parameters. Global parameters affect all vehicles, of any type, when driving in any location on the network. Local parameters are relative to a specific section. They influence all the vehicles while driving in that section. Vehicle type parameters affect all vehicles from one vehicle type (car, truck, bus,…), that circulate in any location of the network (TSS, 2012).

From the vast number of calibration parameters found in AIMSUN (more than fifty), a selection of micro-models parameters to study was made taking into consideration their relevance in the calibration of a traffic signal control complex intersection like the selected as case study (TSS, 2012; Punzo & Ciuffo, 2009; Ciuffo et al. 2012). The selected parameters were:

- *Reaction time, rT;*
- *Reaction time at stop, rTs;*
- *Reaction time at traffic light;*
- *Minimum distance between vehicles, mDv;*
- *Maximum acceleration, MxA;*

These parameters are mostly related to the *car-following* model, as the parameters that affect more the *lane changing* and *gap acceptance* models have small influence in the current type of network (TSS, 2012; Barceló, 2002). The range of errors to be introduced was decided based in the literature (AASHTO, 2001; Deward & Olson, 2007; FHWA/TRB, 2012; Austroads, 2007; Bonneson, 1992; ITE, 1992), ensuring the use of realistic values, although using also extreme values in order to facilitate the detection of patterns between input and output errors.

5. Performance indicators selection

To analyze the network the AIMSUN software provides a wide range of performance indicators, with different levels of aggregation. These indicators can be relative to the entire network, to each section, to turning movements at intersections, to paths defined by the modeller, or to each Origin/Destination pair. Among the global indicators (for the entire network), some of the most important and of common use are (TSS, 2012):

- Travel Time: average time a vehicle needs to travel one kilometer inside the network. This is the mean of all the single travel times for every vehicle that has crossed the network;
- Delay Time: average delay time per vehicle per kilometer. It is the difference between the expected travel time under ideal conditions and the observed travel time. It is calculated as the average for all vehicles in a certain time interval.
- Stop Time: average time at standstill per vehicle per kilometer.
- Total Travel Time: total travel time experienced by all the vehicles that have crossed the network.
- Mean Queue Length: average length of the queue in that section, expressed as the number of vehicles per lane.
- Maximum Queue Length: maximum length of the queue in the section during the considered time interval, expressed as number of vehicles per lane.

An adequate performance indicator to analyze the outputs is crucial for accurate observations and conclusions. For this reason, a preliminary evaluation was performed to select the single performance indicator which was to be used in the analyzes. The various performance indicators were compared, using the output results of the single parameter analysis of *reaction time*. It was observed that all the global indicators evolved in a similar way, with the ones representing average unit values showing a very similar behavior. With these results, the average indicators Delay Time and Stop Time were discarded, as they did not seem to add significant information in relation to the Travel Time.

Total Travel Time was deemed inadequate since it would not be able to eliminate from the results the impacts of different demand levels existent in different scenarios.

Queue related indicators, do show different behavior from the (average) Travel Time based ones. However, some tests performed to evaluate their accuracy in explaining the different conditions of the network, have shown that these would not be so important, since cases of blocking back were not at stake.

Finally, the selected performance indicator was the (average) Travel Time.

6. Types of Analyses Performed

Three types of analyses were made with input error introduced: in a single parameter; in two or three parameters simultaneously; and in a single parameter but considering different network load levels.

6.1. Single parameter analyses

The first step to understand the influence of input errors is to study the introduction of errors individually in relevant calibration parameters. The first parameter studied was the *reaction time* (rT). Although this is one of the most studied parameters, it is also one of the most difficult and expensive to collect in the field. To perform this analysis (analysis nº1), an error to the *reaction time* was introduced, varying its value in the range of 0.50sec to 1.50sec, around the default value of 0.75sec with value increments of 0.05sec (see table 1). The *reaction time* was fixed to the value of the *simulation step*, therefore the values of both were the same during the runs.

Important in queue modeling are the parameters *reaction time at stop* and *reaction time at traffic light*. It is recognized the existence of differences between the real values of these two parameters (Bonneson, 1992;

FHWA/TRB, 2012). Nevertheless, it is usually considered a reasonable simplification to consider both parameters with the same values, since the difference fades away as the queue growths. Therefore, all the next references to *reaction time at stop* refer to *reaction time at traffic light* as well. Errors were introduced within a range [0.90; 2.70]sec, with increments of 0.09sec. This is analysis nº2 (see table 1).

Also potentially important in queue modeling is the parameter *minimum distance between vehicles*. Errors were introduced only to the vehicle type *car*, in the range [0.5; 2.0] meters with increments of 0.1 meters, keeping the default value in AIMSUN of the standard deviation, 0.3 meters. This is analysis nº3 (see table 1).

Another parameter that was chosen to be studied was the *maximum acceleration*. The input error, only to the vehicle type *car*, was in a range of [2.0; 4.0]m/s^2, with increments of 0.10 m/s^2, keeping the standard deviation equal to 0.4 m/s^2. This is analysis nº 4 (see table 1).

Table 1. Analyses carried out, with the range of input errors introduced in the parameters.

Analyses	Parameters with input errors				
	SS - Simulation Step (s)	Reaction time, rT (s)	Reaction time at stop, rTs (s)	Min. distance between vehicles, mDv (m)	MxA - Maximum Acceleration (m/s^2)
1	**[0.5 ; 1.5] 0.05**	**[0.5 ; 1.5] 0.05**	{1.35}	x={1.0}	x={3.0}
2	{0.75}	{0.75}	**[0.9 ; 2.7] 0.09**	x={1.0}	x={3.0}
3	{0.75}	{0.75}	{1.35}	**X= [0.5 ; 2.0] 0.1**	x={3.0}
4	{0.75}	{0.75}	{1.35}	x={1.0}	**X= [2.0 ; 4.0] 0.1**
5	**[0.5 ; 1.5] 0.05**	**[0.5 ; 1.5] 0.05**	**[0.9 ; 2.7] 0.09**	x={1.0}	x={3.0}
6	**[0.6 ; 0.9] 0.05**	**[0.6 ; 0.9] 0.05**	{1.35}	**X= [0.7 ; 1.3] 0.1**	x={3.0}
7	**[0.6 ; 0.9] 0.05**	**[0.6 ; 0.9] 0.05**	{1.35}	x={1.0}	**X= [2.7 ; 3.3] 0.1**
8	**{ 0.60 ; 0.70 ; 0.75 ; 0.85 ; 0.90 }**		{1.35}	**x= {0.8 ; 1.0 ; 1.2 ; 1.3}**	**x= {2.7 ; 2.8 ; 3.0 ; 3.2}**
9*	**[0.5 ; 1.5] 0.05**	**[0.5 ; 1.5] 0.05**	{1.35}	x={1.0}	x={3.0}

**Saturation levels: 80%, 90%, 100%, 110%, 120%*

6.2. Analyses with combination of input error in parameters

In the study of isolated parameters the introduction of input errors in an isolated and independent way was assumed. However, that tends not to be what happens in real situations, and actually, it can be expected that there exists some interdependence between the different errors' impacts (Saltelli et al.,2004). There is, thus, a need to evaluate the potential impact resulting from the simultaneous occurrence of input errors in different parameters. In question there might be potentiation or mitigation effects, resulting from inter-dependencies between effects of different input errors (Ciuffo et al., 2012).

The first analysis with combination of input errors was the variation of the *reaction time* and *reaction time at stop* (incorporating also the *reaction time at traffic light*) simultaneously (analysis nº5), and not separately, like in the single parameters analyses. Since reaction times, in any situation, are connected to the psychomotor capabilities of each driver (Dewar & Olson, 2007), it seems reasonable to expect that a set of drivers that within a certain context show an above the average reaction time, will also show it in different contexts. This has led to the adoption of sets of correlated errors (starting with negative values and finishing with positive ones) in the different types of reaction times as shown in table 1.

It was also considered important to analyze the effect of combining input errors in the parameters *reaction time* and *minimum distance between vehicles*, and combining the effect of input errors in the parameters *reaction time* and *maximum acceleration* as well. This resulted in analyses nº6 and nº7. In these analyses, the range of values tested were restricted to the most realistic ones.

To better evaluate the effects of errors in several input parameters simultaneously, these last three tested parameters were combined in one analysis (analysis nº8). In this case, the range of testing values were also reduced, to limit the computation time, as presented in the table 1.

6.3. Input errors with different network saturation levels

It was considered useful to study the potential of the networks' degree of saturation to influence the level of impact of errors committed in the calibration of input parameters. Input errors were then introduced in a single parameter, but testing its impact level according to the network saturation levels (analysis nº9). In the current study a single parameters analysis to the *reaction time* was performed, comparing the results obtained with saturation levels of 80%, 90%, 100%, 110% and 120% of the initial O/D matrix (see table 1).

6.4. Confidence levels in stochastic modeling

Microsimulation models, namely AIMSUN models, deal with a very large number of stochastic events, leading to different results in each run made in the model. Thus, appropriate measures need to be taken to control the reliability and significance of the results (Antoniou et al., 2014). Generally this problem is addressed by making a sufficient number of runs (replications) of the same situation using different "seeds" to reduce punctual deviatory values (FHWA, 2004; Punzo & Ciuffo, 2009). Hollander & Liu (2008), for example, have revised a number of studies in which the number of runs have gone from 1 to 20. In the present study every result value is based on 30 runs or replications. In total more than 11000 runs were performed to support the results here presented. The number of runs for each analysis is shown in table 2.

Table 2. Number of runs performed to each analysis in total.

Analysis	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Nº of runs	630	630	480	480	630	1470	1470	2400	5*630

7. Results Analyses

7.1. Single parameter analyses

From the analysis with input errors in a single calibration parameter, results can be gathered relating the values of the performance indicator (average) *Travel Time* and the values of the parameters. To analyze the patterns of these relations different regressions were fitted to the results: linear and polynomial of 2nd degree.

In all four parameters analyses, linear and 2nd degree polynomial regressions showed good fitting, slightly better for the polynomial in all cases, having the values of R^2 started at 0.867, except for the *maximum acceleration*, which showed an R^2=0.606. From qualitative residual analyses it was not possible to observe clear residual tendencies, thus suggesting that the regression curves fitted represent the tendencies in a trustable way. It was decided to keep the 2nd degree polynomial based regressions in all these analyses.

Further, to present the results in a comparable and transferable way, it was decided to represent the magnitude of the input parameter as a percentage of the considered "correct" value, and to evaluate the impact over the performance indicator, also in the form of the percentage deviation in relation with the "correct" output. A "correct" value (= *starting value*) would be the value of the parameter found in the field (the real/correct one) and a "correct" output would result from the use in the simulations of the "correct" value for the parameter. The resulting *input error - output error* relations are represented in fig. 2.

Regarding the *reaction time* (*starting value* = 0.75s), it can be seen that for an input error of +20% there is an output error of about +7%, while for an negative input error there is a smaller error in the output. This difference suggests that it is better to have a defect error, than an error by excess. It can also be noticed that the error in the result gets bigger the bigger the "correct" parameter *starting value* is assumed to be.

Regarding the *reaction time at stop*, substantially different patterns can be found, from the ones in the *reaction time* analysis. For the different starting "correct" parameter values the impact in the outputs are quite similar. For an input error of +20% the output error is about +20%, with symmetric relations to negative errors.

The analysis with the *minimum distance between vehicles*, shows an input – output error relation where the output is approximately half of the input error. An input error of +20% results in an output error of 8%, and an input error of -20% results in an output error of -10%. This gives the idea that for this parameter it is better to have an input error by excess than by defect.

Regarding the *maximum acceleration*, an input error of +20% resulted in an output error of -5%, and an input error of -20% resulted in a much more significant output error of 10%.

Globally, in spite of the differences of pattern and magnitude, one could concluded that all the parameter errors have shown a significant impact in the model outputs' quality.

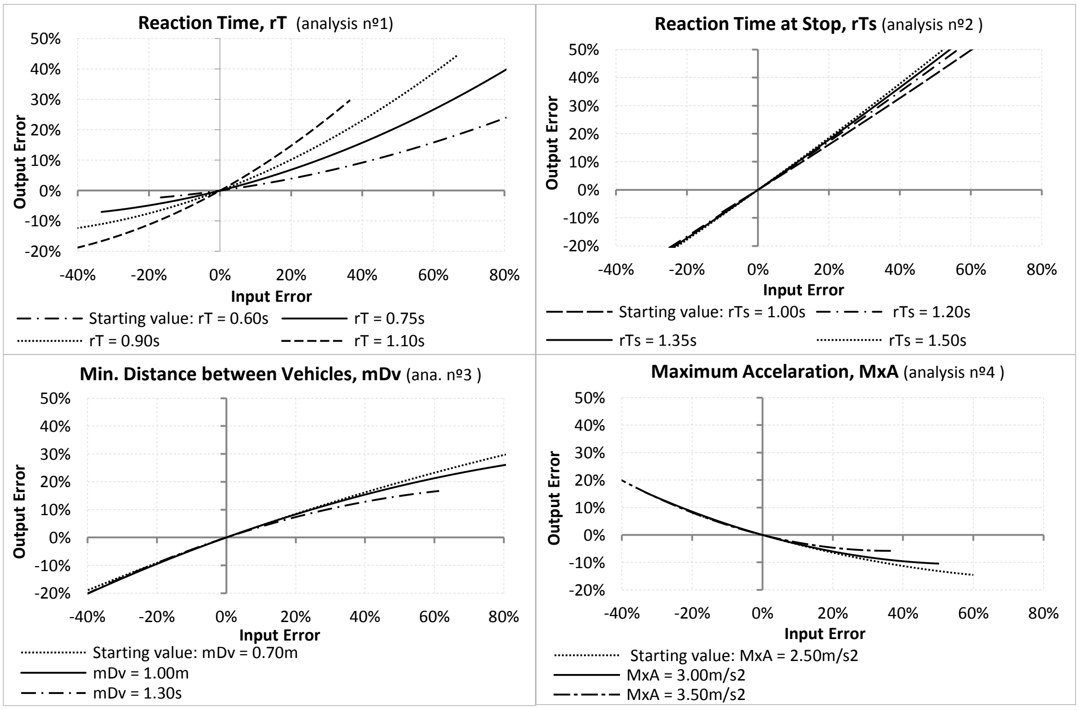


Fig. 2. Relations comparing the size of output error, from the size of input error, with different *starting values*, in single parameters analyses.

7.2. Analyses with combination of input errors in parameters

The first analysis with combination of input errors, considered a combination of *reaction time* and *reaction time at stop*, enabling the comparison of its results against those obtained in the single analyses for both parameters, testing the degree of additivity of the effects. Results, presented in the figure 3, show that, for the most common input error interval (-20%, +20%), the effect of the combination of errors is very similar to the addition of the effects of each error, when considered in isolation. From this, it can be concluded with some confidence that, in this interval, these two types of input errors show independence of impacts, and thus, when combined, result in additive effects.

For the analysis with combination of input errors in the parameters *reaction time (rT)*, *minimum distance between vehicles (mDv)* and *maximum acceleration (MxA)*, the results are presented in figure 4.

Analyzing the results from the combined *rT + mDv* (Fig. 4, (a)) it can be noticed that the *Travel Time* tend to increase as the value for *mDv* increases, for all *rT* values. A linear multiple regression model was tested (1):

$$Travel\ Time = b_0 + b_1.rT + b_2.mDv \qquad (1)$$

The results show a very good fitting (R^2=0.812), with both parameters showing significant influence in the travel time explanation (b'_1=0.783, b'_2=0.446), with a bigger influence attributable to *mDv*.

However, the graphic also shows some not expected irregularities related to the *rT* values. For all the values of *mDv* the *Travel Time* decreases from the *rT*=0.60s till *rT*=0.70s, when it would be expected an increase. Also, the *Travel Time* for the *rT*=0.80s is always smaller than for the *rT*=0.75s. These patterns were not expected, and although they can result from internal processes intrinsic to the software, an inspection to the random seeds generation process was not conclusive.

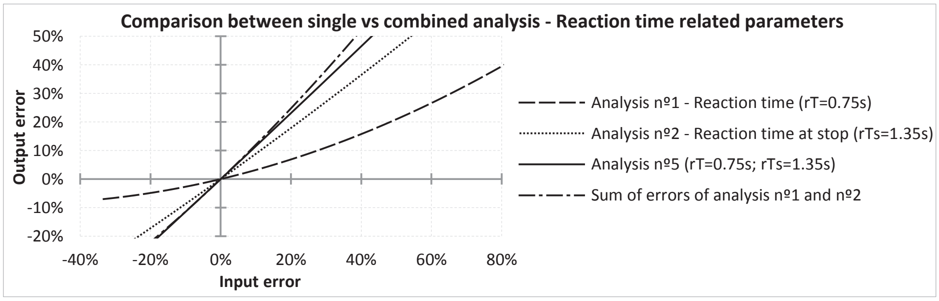


Fig. 3. Comparison of output errors caused by single and combined errors at *reaction time* and *reaction time at stop*.

The results from the combined *rT* + *MxA* (Fig. 4, (b)) analysis*,* based on a linear multiple regression, show a good fitting (R^2=0.581). Further, the parameter *maximum acceleration* has shown a small influence in the results (b'_2=0.225), compared with the influence from the *reaction time* ($b'_1=0.783$). However, as in the previous analysis, in these results it can also be noticed the occurrence of some "strange" irregularities in the *rT*=0.80s and *rT*=0.70s. This phenomenon needs further investigation(*b' coefficient is the b coefficient normalized*). Residual analyses were carried in both analyses, which showed that the regression models produced non skewed and significant results.

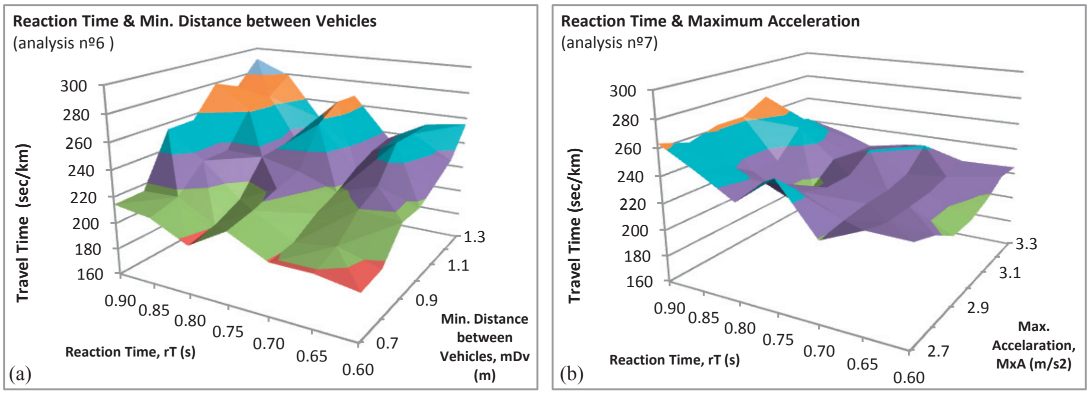


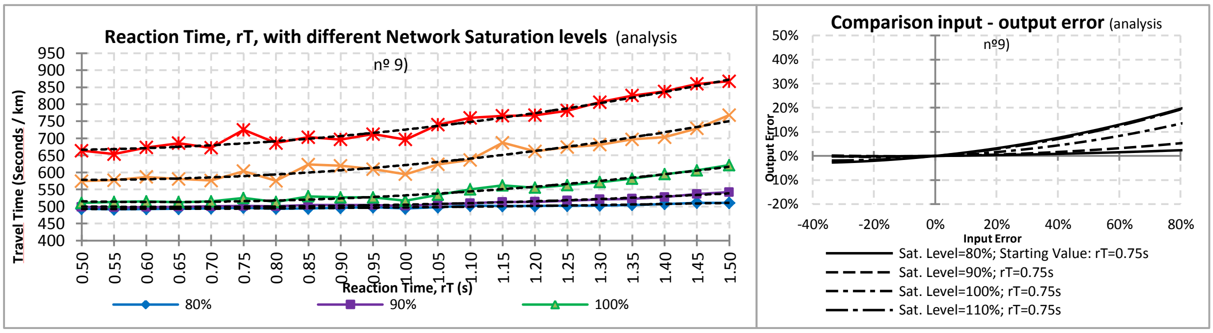
Fig. 4. Results of combined input errors analysis. (a) *reaction time+min. distance between vehicles;* (b) *reaction time+max. acceleration*.

Finally, the analysis of the combination of *rT* + *mDv* + *MxA* parameters (analysis nº8), came to confirm the results from the last two analyses: bigger influence of *mDv*, some influence of *rT* and almost no influence of *MxA*.

7.3. Single parameter analysis with different network saturation levels

In this case, it was tested the output error caused by input error in the *reaction time*, but considering different network saturation levels. The resuls are presented in figure 5. To guaranty that the results were not influenced by the occurrence of fictitious "virtual queues", the road link of the most congested entrance of the intersection was extended, so the results in this analysis cannot be directly compared to the other analysis.

Observing the results, it can be noticed that the ones with input errors applied to a low saturation level network (80% and 90%) are similar (fig. 5 (a)), showing that low congestion levels lead to smaller, and less dependent on the level of demand, impacts of input errors in the model outputs (fig.5 (b)), and tend to have no impact with saturation levels under 80%. On the other side higher congestion levels (110%, 120%) lead to bigger impacts of input errors in the model outputs.



(a) Direct results, with mean virtual queue values. (b) Relation between output error with input error.

Fig. 5. Results of the input error analysis with different network saturation levels (analysis nº9).

8. Conclusions

The aim of the present work was to give a contribution to the identification of which calibration parameters errors tend to affect more the simulated results, and what are adequate precision levels to be achieved in a calibration process.

The results obtained by applying isolated errors to a single calibration parameter, although being generally well explained by regression analyses, showed a considerable diversity of results. For some parameters (*reaction time*, *maximum acceleration*) the input-output errors relations have suggested polynomial patterns for the relation between input and output errors, while others (*reaction time at stop*, *minimum distance between vehicles*) suggested linear patterns. Further, within the most common errors range (-20%, +20%), the different input errors have shown quite different impact potential. For the *reaction time*, this relation between input and output errors is between 4:3 and 4:4, while for the other three parameters this relation is between 4:1 and 4:2, showing the importance of the *reaction time* parameter. It was also possible to conclude that for different reference *"correct" values*, the exact patterns of these relations can be different.

In the analyses where a combination of errors was applied to more than one parameter simultaneously, an almost additive effect of the individual input errors was observed, suggesting a nearly independence of effects.

Furthermore, the impact caused by input errors applied in the *reaction time* and the *reaction time at stop* seem to have approximately the same level of impact.

On the contrary based on the analysis of another combination of input errors it was found that the *maximum acceleration* presented a very small influence, almost negligible, comparing to the *reaction time* and to the *min. distance between vehicles*. Between these last two, it could also be noticed that, for the case study under analysis, the *min. distance between vehicles* had a more significant impact than the *reaction time*.

In the last analysis, where an input error into a single parameter was applied under different network congestion levels, it could be concluded that, generally, the level of impact of input errors in the quality of the outputs is quite positively correlated with the intersection saturation levels. In the case studied with low saturation, represented by 80% to 90% levels, the impact levels tended to increase only modestly, but as the saturation levels grew, the errors' impact tended to grow quite quickly, but only until a certain level of saturation, as expected. For the studied case, after 110% of saturation in the network, the impact of input errors did not increase more (the impact of 120% was the same than that of 110%).

Finally, further investigation needs to be performed, in order to, systematically, study the most important calibration parameters, when applied in different types of networks, working under different congestion levels. As an example, intersection/junctions without regulation by traffic lights should be analyzed, to study the *gap-acceptance* related parameters.

Acknowledgements

This work has been carried out in the framework of projects EMSURE - Energy and Mobility for Sustainable Regions (CENTRO-07-0224-FEDER-002004) and AROUND (FCT project PTDC/SEN-TRA/122114/2010).

References

AASHTO. (2001). A Policy on Geometric Design of Highways and Streets, 4th Edition,. Washington, D.C., USA: American Association of State Highway and Transportation Officials (AASHTO).

Antoniou, C., Barcelo, J., Brackstone, M., Celikoglu, H.B., Ciuffo, B., Punzo, V., Sykes, P., Toledo, T., Vortisch, P., Wagner, P. (2014). Traffic Simulation: Case for guidelines. Luxembourg: Joint Research Centre - European Commission.Austroads. (2006). The use and application of microsimulation traffic models - Austroads Research Report AP-R286/06. Sydney NSW, Australia: Austroads Incorporated.

Austroads. (2007). Microsimulation Standards - Technical Note 2. Victória, Australia: ARRB Group Ltd.

Barceló, J. (2002). Dynamic Network Simulation With AIMSUN. International Symposium on Transport Simulation. Yokoama.

Bonneson, J. A. (1992). Modeling Queued Driver Behaviour at Signalized Junctions - Transportation Research Record 1365. Washington, D.C., USA: Transportation Research Board.

Ciuffo, B., Miola, A., Punzo, V., & Sala, S. (2012). Dealing with uncertainty in sustainability assessment. Luxembourg: Joint Research Centre - European Commission.

Dewar, R., & Olson, P. (2007). Human Factors in Traffic Safety. Tucson, USA: Lawyers & Judges.

FHWA. (2004). Traffic Analysis Toolbox Volume III - Guidelines for Applying Traffic Microsimulation Modeling Software. McLean, VA, United States: FHWA - Federal Highway Administration.

FHWA/TRB. (2012). Revised Monograph on Traffic Flow Theory (1975-2012). Washington, D.C., USA: FHWA - Federal Highway Administration.

Hollander, Y., & Liu, R. (2008). The principles of calibrating traffic microsimulation models. Transportation - Springer, Volume 35, Issue 3, pp 347-362.

ITE. (1992) Traffic Engineering Handbook. Washington, DC., USA: Institute of Transportation Engineers (ITE).

Punzo, V., & Ciuffo, B. (2009). How Parameters of Microscopic Traffic Flow Models Relate to Traffic Dynamics. Transportation Research Record: *Journal of the Transportation Research Board*, 249–256.

Saltelli, A., Tarantola, S., Campolongo, F., Ratto,M., (2004). Sensitivity Analysis in Practice. A guide to Assessing Scientific Models. Wiley

Sargent, R. (2000): “Verification validation, and accreditation of simulation models”. Proceedings of the 2000 Winter Simulation Conference. J. A. Joines, R. R. Barton, K. Kang, and P. A. Fishwick, eds.

TSS. (2012). Aimsun 7 Dynamic Simulators User's Manual. TSS – Transporting Simulation Systems.

Vasconcelos, L., Seco, Á, Silva, A. (2009), Accessing the Impact of Quantification Errors in Traffic Simulation Models, Proceedings of the 2009 IASTED International Conference on Modelling, Simulation and Optimatization, Palma de Mallorca, Spain